

Saccadic and perceptual performance in visual search tasks.

I. Contrast detection and discrimination

Brent R. Beutter

MS 262-2, NASA Ames Research Center, Moffett Field, California 94035-1000

Miguel P. Eckstein

Department of Psychology, University of California, Santa Barbara, California 93106-9660

Leland S. Stone

MS 262-2, NASA Ames Research Center, Moffett Field, California 94035-1000

Received July 29, 2002; revised manuscript received December 4, 2002; accepted December 16, 2002

Humans use saccadic eye movements when they search for visual targets. We investigated the relationship between the visual processing used by saccades and perception during search by comparing saccadic and perceptual decisions under conditions in which each had access to equal visual information. We measured the accuracy of perceptual judgments and of the first search saccade over a wide range of target saliences [signal-to-noise ratios (SNRs)] in both a contrast-detection and a contrast-discrimination task. We found that saccadic and perceptual performances (1) were similar across SNRs, (2) showed similar task-dependent differences, and (3) were well described by a model based on signal detection theory that explicitly includes observer uncertainty [M. P. Eckstein *et al.*, *J. Opt. Soc. Am. A* **14**, 2406 (1997)]. Our results demonstrate that the accuracy of the first saccade provides much information about the observer's perceptual state at the time of the saccadic decision and provide evidence that saccades and perception use similar visual processing mechanisms for contrast detection and discrimination.

OCIS codes: 330.1880, 330.2210, 330.4060, 330.4300.

1. INTRODUCTION

In everyday life, if you are not sure what a person is thinking or doing, you can get a better understanding by watching the person's eyes to determine what they are looking at. The goal of this study is to determine whether similar visual processes mediate saccades and perception and how much can be learned about visual perception by examining saccadic eye movements. Saccades are fast, ballistic eye movements used to rapidly change gaze position from one region in the visual field to another. To explore and gather information about their environment, humans make frequent saccades to point their high-resolution foveae at the object or the spatial location of current interest. Clearly, saccades and perception are related; humans can deliberately make a saccade to foveate a chosen object to gain more information about it, yet most of our frequent saccades occur without such conscious deliberation. Each saccade is the result of a neural decision, which specifies where the saccade should land and when to initiate it. We are interested in how the brain makes such decisions and the relationship between this motor decision and perception when the location of the target is not completely obvious or predetermined. Do saccades and perception share visual processing about the possible target location, or are they controlled by separate neural pathways that make independent decisions based on different information?

Saccadic eye movements are especially important in performing visual search tasks in which the observer must scan a display to locate a target. While a number of studies have examined saccadic and perceptual performance in search and searchlike tasks, few have directly addressed the question of whether or not saccades and perception use the same neural estimate of target location. If the saccadic and perceptual decisions are mediated by the same neural processing, then saccadic and perceptual performance should be similar across a wide range of experimental conditions (e.g., target-distractor similarity, type of task, etc.). However, to pose this question in a meaningful way, it is necessary to ensure that both saccades and perception have access to the same visual input. In particular, processing times and target eccentricities must be matched. It is also essential that a single performance metric be used for both saccades and perception. In this study, humans were asked to search a visual display to find a target. We recorded the saccades that observers used to search the display and their perceptual judgments of where the target was located. We designed experiments such that saccades and perception would have access to the same visual information and used signal detection theory (SDT) to quantify and compare the accuracy of the perceptual and saccadic decisions under matched conditions.

A. Comparing the Information Mediating Saccadic and Perceptual Decisions

Our first goal is to compare saccadic and perceptual performance under conditions in which the visual information available to both is approximately the same. As in our earlier studies^{1,2} (see also Findlay³), we chose decision accuracy as our metric and compared the accuracy of the perceptual decision after a brief presentation with that of the first saccadic decision in an extended search. Our observers searched a visual display with ten possible target locations and reported their ten-alternative forced-choice (10-AFC) perceptual decision. We used the eye movement data to assign a 10-AFC saccadic decision to each saccade by choosing the element location closest to the saccadic end point. Thus, for both saccades and perception, we measure the accuracy of a 10-AFC decision. To match the processing times for saccades and perception, we measured performance in two separate conditions. In both the eye movement (EM) and fixation (FIX) conditions, the observer's task was simply to find the target and report its location. In the EM condition, observers were free to use saccades to search the display, and we measured the 10-AFC accuracy of the first saccadic decision and the saccadic latency. In the FIX condition, to restrict perception to have approximately the same amount of time available to process the visual image as the first saccade, we used a brief display (followed by a high-contrast mask) and observers maintained central fixation (there was no time to make a saccade). Central fixation ensured that the eccentricity of the element locations was matched for the saccadic and perceptual decisions. The high-contrast mask limited the time available for perceptual processing to the stimulus presentation time. The stimulus duration was chosen so that perceptual and saccadic processing times were approximately equal (see Section 2). Although we recorded the perceptual responses for both the FIX and EM conditions, we are mainly interested in perception for the FIX condition because its perceptual processing time is nearly matched to the saccadic processing time. We do not compare saccadic performance with EM perception (free viewing), because this perceptual decision has additional information based on up to 4 s of processing time and multiple fixations at different image locations. We report the EM perception data separately because they may be of interest to some readers.

B. Measuring How Saccadic and Perceptual Decision Accuracies Depend on Salience

Our second goal is to measure how saccadic and perceptual decisions during a search depend on the salience of the visual target relative to the distractors. To quantify "salience," we use stimuli with a wide range of signal-to-noise ratios (SNRs) and use SDT to compare the measured perceptual and saccadic performances with each other¹ and with the performance of the ideal observer.^{4,5} To measure performance under a wide range of difficulties, we choose five SNRs such that perceptual performance for the most difficult condition was close to chance while that for the easiest was close to perfect. Numerous other studies have attempted to determine if and how saccadic eye movements during a visual search task depend

on the visual information in the stimulus image. Many researchers have found that saccades are guided by visual information,^{3,6-10} while others have suggested that, under some conditions, they are not.¹¹⁻¹³ We believe that one of the reasons for the disparate results is that the various studies used displays with very different target saliences. Hooge and Erkelens⁹ have shown that changes in salience or discriminability can affect saccadic target selection and fixation duration. We use a simple disk stimulus with added external noise, which allows us to objectively quantify salience as well as its effect on performance.

C. Measuring How Saccadic and Perceptual Decision Accuracies Depend on Task

Our third goal is to determine how the type of search task influences saccadic and perceptual performance. To investigate this, we examined saccadic and perceptual performance in two different search tasks: a contrast-detection task in which observers searched for a single dim target disk embedded in a noisy background and a contrast-discrimination task in which observers searched for the highest-contrast target among ten suprathreshold disks, again embedded in noise. Previous studies^{14,15} found qualitative differences in perceptual performance between contrast-discrimination and contrast-detection tasks. It is, however, unknown if such a difference is also evident in saccadic performance. If saccades and perception share visual processing, then saccadic performance should have similar task-dependent differences. On the other hand, if saccadic decisions are controlled by a different type of neural processing, for instance a more rudimentary/elementary processing, which is mostly sensitive to the presence or the absence of an object and less able to discriminate differences between objects, then it is possible that the task type will affect saccadic and perceptual performance differently.

D. Modeling Saccadic and Perceptual Performance with Signal Detection Theory Models

Our fourth goal is to determine if saccadic performance can be well described by the same SDT model that has been previously shown to provide a good explanation of perceptual performance across a wide range of saliences and a wide variety of visual tasks.^{5,16-19} Specifically, we will use a suboptimal Bayesian model, which postulates that observers make decisions by looking at the outputs of noisy neural mechanisms specified by their template (receptive field) and internal noise level, to determine which target location has the highest likelihood of containing the target. Comparison of the model's best-fitting parameters for saccades and perception will allow a succinct and meaningful quantitative comparison of the neural processing underlying spatial localization for saccades and perception.

2. METHODS

A. Stimuli

The stimuli were displayed on a 21-in. Philips Brilliance 21A color monitor using the AT Vista video display system hosted by a 486 personal computer. The monitor was run in noninterlaced 60-Hz refresh-rate mode with 640

$\times 486$ resolution (pixel size 0.59 mm). At the 57-cm viewing distance, the full display subtended $38^\circ \times 29^\circ$. Luminances were measured by using a PhotoResearch 880 photometer and linearized by using a lookup table. The stimuli were constructed by adding an $18^\circ \times 18^\circ$ noise pattern to a uniform gray background (31.6 cd/m^2 , which corresponded to a lookup table value of 172). For each pixel, the luminance was the sum of the gray background and a Gaussian-distributed, spatially uncorrelated noise sample (rms contrast 26%, power spectral density $2.4 \times 10^{-4} \text{ deg}^2$).

There were ten element locations, which were evenly spaced around a circle of radius 5.9° centered on the fixation point. To minimize the observer's spatial uncertainty, each element (target or distractor) was presented in the center of a 2.4° square outline (0.4 cd/m^2). The target and the distractors were 0.35° diameter, Gaussian-blurred ($\sigma = 3.5$ arc min) disks, which were added to the noise samples. In the contrast-detection experiment [Fig. 1(a)], the target was chosen to be at one of the ten possible locations, and the other element locations merely contained noise samples (equivalently, the distractors had zero contrast). The target disk had five different peak contrast values (13.2%, 19.8%, 26.3%, 32.9%, and 39.5%) corresponding to five SNRs (2.0, 2.9, 4.2, 5.2, and 6.3, respectively). In the contrast-discrimination experiment, one location contained the target, and the nine other locations contained distractors. To minimize the effective

positional uncertainty, we chose the distractor (pedestal) contrast to be large enough so that it was easily detectable ($d' > 4$). The peak contrast of each distractor was 52.7%, and the target's peak contrast was incremented by values similar to those used in the contrast-detection experiment, thereby producing five contrasts (65.9%, 72.5%, 79.0%, 85.6%, and 92.2%) corresponding to five SNRs (2.1, 3.1, 4.2, 5.2, and 6.3, respectively).²⁰ The maximum contrast that we could display was 130% (72.5 cd/m^2), which is high enough to ensure that the image luminance values were rarely clipped (generally, $<1\%$ of pixels with a worse case of 7%).

B. Experimental Procedures

Three observers (two authors, one naive) with normal vision participated in this experiment. They were asked to search the display and find the target (the only disk in the detection task and the brightest disk in the discrimination task) while viewing the stimulus binocularly in a dimly lit room. On each trial, their eye movements and perceptual decision were recorded. Head movement was minimized by the use of a bite bar. Observers were required to fixate a central $1.2^\circ \times 1.2^\circ$ fixation cross before each trial began. They then initiated the trial with a button press. After a delay of 500 ms, one of the search images was presented for either a brief duration (150 ms) or a long duration (4 s). Immediately following the presentation of the search image, a high-contrast noise field (spatially uncorrelated noise with an rms contrast of 74% and a power spectral density of $1.94 \times 10^{-3} \text{ deg}^2$) was presented for 200 ms. Finally, a uniform gray (31.6 cd/m^2) response image, which contained the box outlines and a rotatable arrow, was presented. At the end of each trial, observers indicated their 10-AFC perceptual decision by pointing the arrow at the element location that they thought had contained the target, and then they pressed a mouse button to record their decision. We did not provide trial-by-trial feedback to our observers because we did not want to alter their natural oculomotor behavior or, worse yet, differentially shape oculomotor and perceptual behavior.

For both the detection and discrimination tasks, there were two different stimulus durations that defined the EM and FIX conditions. In the EM condition, observers were given up to 4 s to search the display and find the target. In most earlier studies of saccades in visual search, the observer's task was to make a saccade to the target as quickly as possible. To avoid biasing the observers' search strategies and saccades, we chose not to give an explicit instruction to saccade to the target or any other eye movement instruction. The long duration used in this condition gave the observers time to make many saccades while the stimulus was still present. In the FIX condition, the stimulus duration was 150 ms. As this duration was less than the normal saccade latency, observers maintained fixation on the central cross for the whole presentation. The 150-ms duration was chosen so that the processing time available to perception was approximately matched to that available to the first saccades in the EM condition. Pilot studies suggested that saccadic latencies were ~ 240 ms, which corresponds to a saccadic

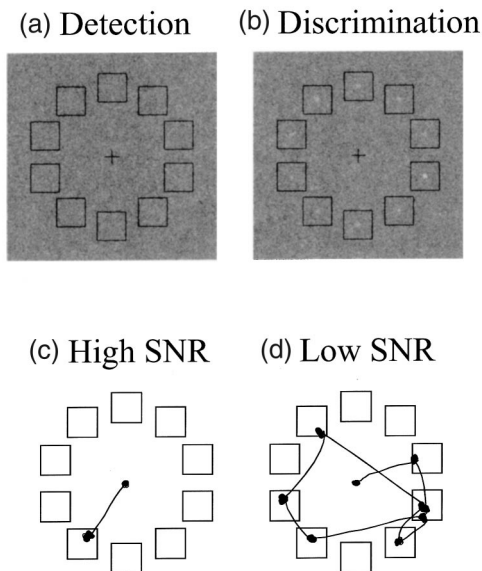


Fig. 1. (a) Example of a stimulus image for the contrast-detection experiment. The nontarget boxes contain only noise, while the target box has a low-contrast disk added to the noise. (b) Example of a stimulus image for the contrast-discrimination experiments. The nontarget boxes each contain distractors of equal contrast, while the target box has a higher-contrast disk. (c) Eye movement data from an "easy" trial (SNR = 6.3) in the contrast-detection experiment for observer JL. This trial contained a single saccade to the target box, so the first saccadic decision was correct. (d) Eye movement data from a "difficult" trial (SNR = 2.0) in the contrast-detection experiment for observer JL. This trial contained eight saccades. The first saccade was to a distractor box, so the first saccadic decision was incorrect. The second, sixth, and last saccades were to the target box.

processing time of ~ 150 ms, given the assumption of ~ 90 ms of delay between the decision to make a saccade and its actual initiation.^{21,22}

Each day, observers ran four blocks (EM discrimination, EM detection, FIX discrimination, and FIX detection) in a random order. Each block consisted of 20 trials at each of the five SNRs for a total of 100 randomly interleaved trials. Each block was repeated 15 times, producing a total of 300 trials for each SNR (three repetitions of 100 different noise samples), each task, and each EM condition. We reduced the effects of cognitive expectation²³ by using stimuli in which the target appeared in one of ten known locations with equal probability and by randomly interleaving all SNR conditions. We found similar results in a previous experiment in which we blocked SNRs.²⁴ Observers ran two practice days before gathering the data presented.

C. Eye Tracking

The position of the left eye was measured with an infrared (IR) video-based eye tracker (ISCAN Inc., custom-built for NASA) sampling at 240 Hz, hardware synchronized to our display monitor. An IR light source illuminated the observer's left eye, which viewed the stimulus through a mirror that transmitted visible but reflected IR light. The tracker computed the horizontal and vertical positions of the pupil in uncalibrated eye-tracker coordinates. Both before and after each run, we calibrated the eye tracker by having observers fixate a series of nine crosses arranged in a $12^\circ \times 12^\circ$ grid.²⁵ The crosses were presented in a fixed pseudorandom order, and each was shown at least twice. For each fixation, the mean eye position and its standard deviation were calculated. The calibrated eye positions were computed as linear functions of the tracker outputs. Two sets of three parameters were used to compute the horizontal and vertical positions, respectively: an offset, a gain, and a cross-talk term. The calibration parameters were determined by optimally fitting the mean eye-tracker outputs to the known locations of the fixation points. The calibration data were well fitted by these six linear parameters; the reduced χ^2 ranged from ~ 0.5 to ~ 1.5 . The eye-tracker positional precision was generally better than 0.15° , estimated from the mean standard deviation of eye position during a fixation.

Despite the use of a bite bar, observers periodically made small changes in their head position or orientation. This generated small overall offsets in the position of the pupil. Because we were interested in determining the absolute gaze position of the observer (i.e., relative to the earth-bound display), we needed to correct for changes in the offset parameters caused by the possible small head movements between trials. Thus, for each trial, we rezeroed the offset parameters based on the initial fixation before each stimulus presentation. Inspection of the scan paths for many trials verified that this approach did a good job of correcting for occasional small head movements.

D. Saccade Detection and Fixation Analysis

We used a low-pass filtered differentiator (-3 -db cutoff at 42 Hz) to detect saccades. The total eye velocity (the

square root of the sum of the squares of the horizontal and vertical velocities) was compared with a threshold, which was set to detect saccades larger than 0.7° . Blinks were also detected and eliminated. The eye-position data were analyzed to obtain a sequence of fixations separated by saccades. For each fixation, we computed the mean fixation location by averaging the position data from the end of the prior saccade to the start of the subsequent saccade. For the FIX conditions, we discarded trials in which central fixation was not maintained throughout the trial (1.7° window). For the EM conditions, we discarded trials with anticipatory saccades (outside of the 1.7° fixation window with latency < 90 ms). Few trials ($\sim 2\%$) were discarded.

To compare the saccadic decisions with the perceptual decisions, we converted the first saccade (outside of the 1.7° fixation window) in the EM conditions into a 10-AFC choice. We defined the saccadic decision to be the element closest to the saccade's end point location.^{1,3} This method does not penalize hypometric saccades generated in the correct direction. However, choosing an alternative, stricter decision criterion does not significantly affect the results.¹ Inspection of the data showed that indeed most saccades were directed unambiguously toward a specific element location, i.e., few saccades were made to regions near the decision boundaries between element locations. Examples of typical EM traces are shown in Figs. 1(c) and 1(d). In high-SNR trials, observers typically made few saccades [Fig. 1(c)], while in the more difficult, low-SNR trials, observers made many saccades [Fig. 1(d)]. While the entire saccadic scan path is potentially interesting, because we are concerned with comparing saccadic and perceptual performance under approximately matched temporal and spatial conditions, we limit our analysis in this paper to the first saccadic decision. A limitation of our analysis is that it thresholds the landing position of the first saccade and neglects possible information contained in the dynamics of the saccade or its exact landing position. As has been recently noted^{26,27} many search saccades are curved. Some of the saccades made by our observers also appear to follow curved trajectories and may contain additional information about saccadic processing. We restrict our saccadic analysis to an examination of the 10-AFC saccadic decision, so that it is commensurate with the 10-AFC perceptual decision.

E. Data Analysis

Because our tasks required finding a target signal in the presence of noise, we used SDT to analyze the data. In particular, this approach allows us to compare saccadic and perceptual performance with that of the ideal observer and to quantitatively measure task-dependent performance differences. To do this, we transformed percent correct into d' , an SDT measure of performance,⁴ and plotted performance as a function of SNR. For an N -AFC task, the relationship between percent correct (PC) and d' is described by

$$\text{PC}(d', N) = 100(2\pi)^{-1/2} \int_{-\infty}^{+\infty} dx \exp\left[\frac{-(x - d')^2}{2}\right] \times [\text{erf}(x)]^{N-1}. \quad (1)$$

F. Model Calculations

We begin by describing a simple SDT model without uncertainty.^{4,5,16,19} We then describe an SDT model that explicitly includes effects of intrinsic uncertainty (see Appendix A for mathematical details of the models). The simple model (with zero uncertainty) assumes that the observer makes a saccadic/perceptual decision by monitoring the output of visual mechanisms and choosing the location that has the mechanism with the maximal response. The observer monitors exactly one mechanism corresponding to each possible target location (a total of ten mechanisms). We assume that the response of each mechanism is a Gaussian-distributed variable with equal standard deviation σ (corresponding to the sum of internal and external noise) and mean μ , which is linearly proportional to the element contrast. Thus all the mechanisms responding to distractors have responses with equal means μ_D , while the mechanism responding to the signal has a higher mean, μ_S . For this simple model, performance depends only on the ratio of the difference between the signal and distractor means to their standard deviation. This quantity is defined to be d' and is linearly related to the SNR:

$$d' = \frac{(u_S - u_D)}{\sigma} = \alpha \times \text{SNR}. \quad (2)$$

The relationship between d' and percent correct is described by Eq. (1). The single free parameter, the slope α , measures how good performance is relative to that of the ideal observer. The best possible performance, that of the ideal observer, is achieved when α is equal to unity. Mismatches between the template that the mechanisms use and the signal's spatial profile, as well as internal noise, produce values of α that are less than unity. This simple model clearly cannot explain our detection data, because it predicts that the d' data lie on a line that passes through the origin.

The SDT model with uncertainty²⁸ has, in addition to the slope parameter α , an additional free parameter, the uncertainty number U .^{16,19,30} To explicitly incorporate observer intrinsic uncertainty, this model assumes that for each of the ten possible target locations, observers monitor one mechanism matched to the target location and U additional irrelevant mechanisms (orthogonal to the signal template), which are also orthogonal to each other (independent responses). Again, the responses of all the mechanisms have the same standard deviation. The mean response of the irrelevant mechanisms is zero because they are orthogonal to the signal and the distractors. The mean response of the mechanisms responding to the distractors/target is proportional to the contrast of the distractors/target, and the difference between the target and distractor means is proportional to the SNR. As above, a decision stage monitors the output of all of the mechanisms and chooses the location corresponding to the mechanism with the largest response.³¹ This model correctly chooses the signal location if the mechanism responding to the signal or any of the irrelevant mechanisms corresponding to the signal location has the highest response, and it is incorrect if any of the other responses is highest.

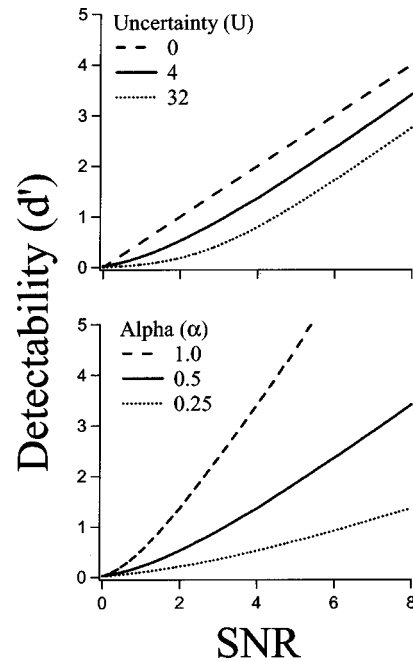


Fig. 2. Effects of the two fitted parameters, alpha (α) and uncertainty (U), on the model predictions. The top graph shows the effect of U on detectability for $\alpha = 0.5$. The bottom graph shows the effect of α on detectability for $U = 4$.

To illustrate the effects of the two fitted parameters, slope (α) and uncertainty (U), plots of the model predictions are shown in Fig. 2. The top graph of Fig. 2 plots performance for three different U values with α fixed at 0.5. For $U = 0$, there is no uncertainty and the detectability is a linear function of SNR with slope equal to α . As U increases, overall performance decreases and the plots become curved for low SNRs and are approximately straight lines for higher SNRs. The bottom graph of Fig. 2 plots performance for three different α values with U fixed at 4. Performance is best for high values of α and decreases as α decreases. Changing α is equivalent to a linear rescaling of the x axis.

3. RESULTS

A. Matching Visual Processing for Saccadic and Perceptual Decisions

The brief display duration in the FIX condition was chosen so that the perceptual processing times for this condition were similar to the saccadic processing times in the EM condition; yet we needed to measure the actual saccadic latencies to determine how well the two were matched. The retinal eccentricities of the element locations were identical for FIX perception and EM saccades because in the FIX condition central fixation was required throughout the trial, and, for EM saccades, the initial saccade was also based on central fixation.

The visual processing time available to make a saccadic decision is shorter than the corresponding saccadic latency because there is a delay between the time the decision to execute a saccade is made and the time saccadic

execution begins. We assumed that this delay was ~ 90 ms based on previous empirical results.^{21,22} Thus, for processing times to be equal, the saccadic latencies should be 90 ms longer than the 150-ms display duration in the FIX condition, and latencies should be similar across SNR and task. Figure 3 shows the median latencies of the first saccade in the EM condition plotted as a function of SNR for the detection and discrimination tasks. For all three observers, the saccadic latencies did not depend strongly on task or SNR. The perceptual and saccadic processing times were well matched for two of the observers, BB and JL. Their overall median saccadic latencies (BB, 229 ms; JL, 242 ms) were ~ 86 ms longer than the FIX display duration. The third observer, LS, had shorter saccadic latencies (median of 158 ms), and thus a saccadic processing time of ~ 68 ms, which was much shorter than the 150-ms perceptual processing time.

Saccadic latencies for all three observers decreased by ~ 25 ms across the threefold increase in SNR. Thus the increased saccadic decision accuracy observed at higher SNRs (see below) cannot be caused by a speed-accuracy trade-off. The weakness of the dependence of latency on SNR is not surprising, as all of the SNRs were randomly interleaved; i.e., before each trial, observers were unaware which SNR would be presented, so they could not adjust their strategy. Latencies for the discrimination task (solid symbols) were on average ~ 15 ms longer than those for the detection task (open symbols). This shows that the task-dependent performance differences described below are also not due to a speed-accuracy trade-off.

Figure 4 shows histograms of the latency distributions (accumulated over SNR) for both correct and incorrect saccadic decisions in both the detection and discrimination tasks. For all observers, in both tasks, the distributions were unimodal and similar for the incorrect and correct decisions. Correct trials tended to have slightly shorter latencies (~ 10 ms) than incorrect trials. This is also inconsistent with a speed-accuracy trade-off and provides strong evidence that the same decision strategy was used on both correct and incorrect trials.

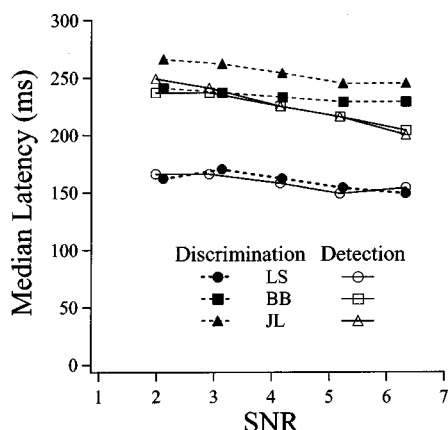


Fig. 3. Median saccadic latencies (EM condition) as a function of SNR for the detection (open symbols) and discrimination (solid symbols) tasks for each of the three observers. In this figure, the latencies for correct and incorrect saccadic decisions have been combined.

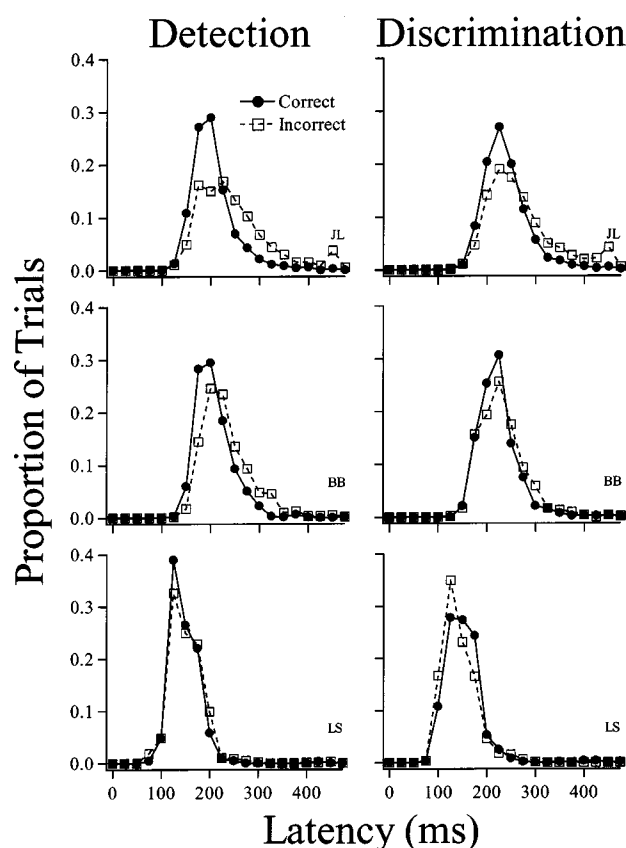


Fig. 4. Histograms of the saccadic latencies (EM condition) for correct (solid circles) and incorrect (open squares) saccadic decisions for each of the three observers (bin size is 25 ms). In this figure, the latencies for the different SNRs have been combined.

B. Saccadic and Perceptual Decision Accuracy

In Fig. 5, we compare the raw performance data (proportion correct) on the detection task with those on the discrimination task for naive observer JL. The accuracies of all three decisions (first saccade, perception limited to 150 ms, and perception up to 4 s) show similar task-dependent differences. For low SNRs, performance on the detection task is worse than that on the discrimination task. For increasing SNR, detection performance improves faster than discrimination performance. At the highest SNRs, detection performance is better than or equal to discrimination performance.

To better interpret the performance data, we converted the data to d' units and replotted them in Fig. 6. For all three observers, both the saccadic and perceptual data show similar SNR trends and similar task dependencies. For the discrimination task, the detectability is approximately directly proportional to SNR (a line through the origin). For the detection task, detectability also appears to be a linear function of SNR but to have a higher slope and a negative y intercept. To quantify these task-dependent differences, we fitted the data with an SDT model, described by Eckstein *et al.*¹⁹ The model fits are plotted as solid (FIX perception), dashed (saccades) (Fig. 6), and dotted curves (EM perception).

Our primary interest is comparing the accuracy of the first saccadic decision with that of the perceptual decision in the FIX condition (solid circles and open squares, re-

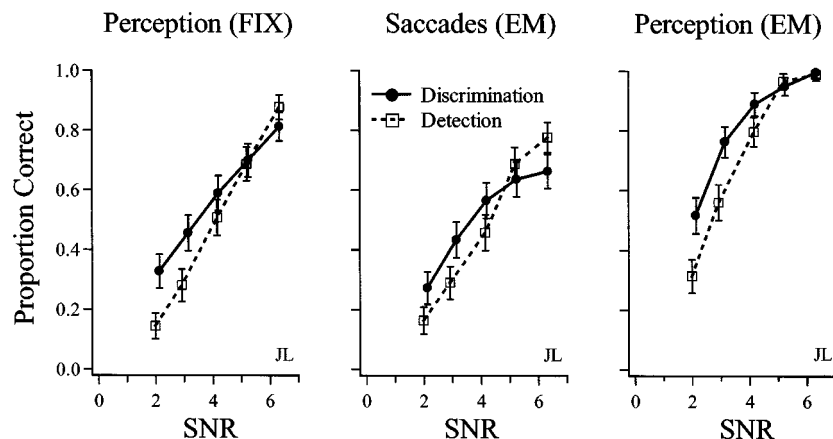


Fig. 5. Proportion of correct decisions for the discrimination (solid circles) and detection (open squares) tasks for observer JL (error bars represent standard errors): (a) accuracy of the perceptual decision in the fixation (FIX) condition, (b) accuracy of the first saccadic decision in the EM condition, (c) accuracy of the final perceptual decision in the EM condition.

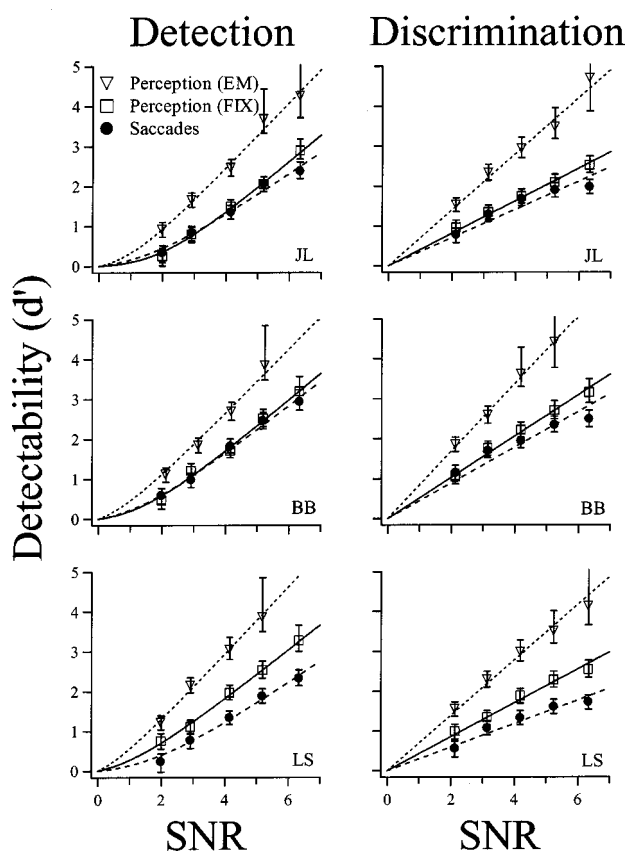


Fig. 6. Accuracy of the FIX and EM perceptual decisions and the first saccadic decision for both the detection and discrimination tasks. The accuracy in d' units is plotted as a function of SNR. The solid circles show the accuracy of the first saccadic decision, the open squares show the perceptual accuracy in the FIX condition, and the open triangles show the perceptual accuracy in the long-duration EM condition. The lines through the points are optimal fits of the signal detection uncertainty model. Error bars show the standard error of the mean.

spectively). For the two observers with approximately matched processing times (JL and BB; see Subsection 3.A), in both the detection task (Fig. 6, left-hand plots) and the discrimination task (Fig. 6, right-hand plots), saccadic and perceptual (FIX) accuracies are very similar.

In the detection task, they are not significantly different ($p > 0.05$, two-way within-observer analysis of variance). In the discrimination task, the overall accuracy of saccades is slightly but significantly ($p < 0.05$) lower (on average, BB is 8% lower and JL is 11% lower) than that of (FIX) perception. The short saccadic latencies of the third observer (LS) (see Subsection 3.A) likely explain why his saccadic accuracy is significantly ($p < 0.001$) lower than his (FIX) perceptual accuracy for all SNRs in both the detection and discrimination tasks. Note, however, that, even for this observer, the task dependencies are similar for perception and saccades.

We also analyzed the accuracy of the perceptual decision in the long-duration EM condition (the open triangles in Fig. 6) and found that it is significantly better than that of the first saccade and that of perception in the short-duration FIX condition at all SNRs. This is not surprising and is the result of at least two important factors.¹ First, in the EM condition, observers can acquire additional stimulus information by foveating potential target locations, while the first saccade and perception in the FIX condition have access only to peripheral information. Second, in the EM condition, perception has up to 4 s of processing time, while the first saccadic and perceptual (FIX) decisions are limited to ~ 150 ms.

C. Efficiency Analysis

To quantify how well humans perform a task, it is useful to measure how much of the available stimulus information they use to make their decision, employing a metric that is independent of task and task difficulty. This is generally done by comparing human performance with the best possible performance (that of the ideal observer) by computing the absolute efficiency,^{5,33} which is defined as the square of the ratio of the human-observer d' to the ideal-observer d' (which is the SNR). Figure 7 plots the absolute efficiency of the first saccade and of perception (FIX) as a function of SNR.

For detection, the two observers with properly matched processing times (BB and JL) had nearly identical absolute efficiencies for saccades and perception, but, for discrimination, their saccadic absolute efficiencies were lower than their perceptual absolute efficiencies (on average, 14% and 21% lower, respectively). For both tasks,

the observer with improperly matched processing times (LS) had saccadic efficiencies that were much lower than his perceptual efficiencies (on average, 58% lower for detection and 52% lower for discrimination).

For all observers, both first saccades and perception showed similar trends as a function of SNR and task. In the discrimination task, absolute efficiencies were nearly constant for both saccades and perception at all SNRs (except for BB, who showed a small decrease in efficiency as SNR increased). In the detection task, absolute efficiencies are low for the lowest SNR, increase as SNR increases, and are similar to those for the discrimination task at the highest SNR.

To compare the amount of information used by first saccades and by perception under approximately matched conditions, we also computed their relative efficiency, defined as the ratio of the squared d' values.¹ For all observers, the relative efficiency of saccades to perception is nearly independent of SNR. For detection, the two observers with properly matched processing times had mean relative efficiencies (BB, 1.03%; JL, 1.06%) that were not significantly different from 100% ($p < 0.05$ t test), while the observer with improperly matched processing times (LS) had a mean relative efficiency that was much lower, 42%. For discrimination, BB and JL had high mean relative efficiencies (86% and 79%, respectively), while again LS had a lower mean relative efficiency (48%).

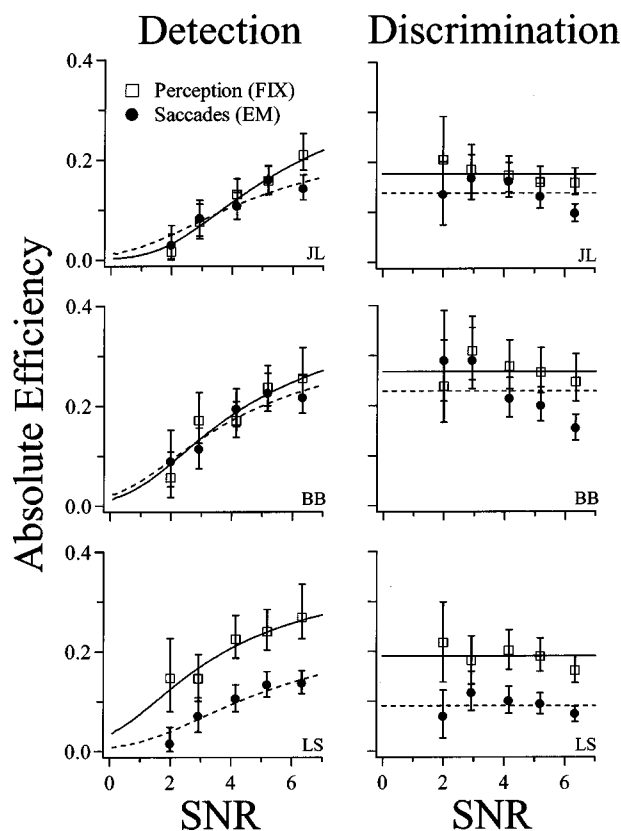


Fig. 7. Absolute efficiencies for (FIX) perception and saccades plotted as a function of SNR for the three observers. The model predictions are shown as the curves [perception (solid), saccades (dashed)]. For the detection task, efficiency increased as SNR increased, while for the discrimination task, efficiency was nearly constant as a function of SNR.

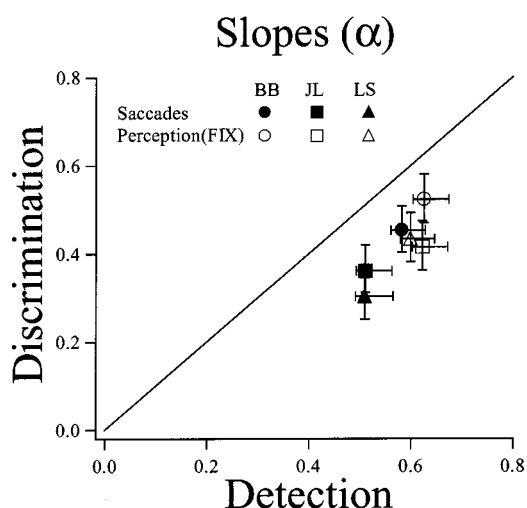


Fig. 8. Maximum-likelihood values [from the fits to Eq. (A7)] of the slope parameter (α) for FIX perception and saccades. Error bars represent 95% confidence intervals. For both saccades (solid symbols) and perception (open symbols), the slopes for detection were higher than those for discrimination.

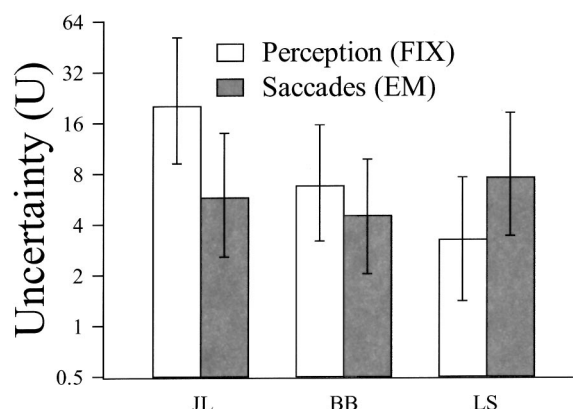


Fig. 9. Maximum-likelihood values [from the fits to Eq. (A7)] of the uncertainty number (U) for the detection task for FIX perception (open bars) and the first saccade (solid bars). Error bars represent 95% confidence intervals. For the discrimination task, the best-fitting uncertainty was zero for all observers for the saccadic data and both sets of perceptual data (data not shown).

We also examined the efficiency of the perceptual decision in the EM condition (not shown). The absolute efficiencies for this condition were consistently higher than those for both the saccadic and (FIX) perceptual conditions, although the task and SNR trends were similar. Again, for all observers, absolute efficiency increased as SNR increased for detection (JL, from 21% to 50%; BB, from 31% to 55%; LS, from 37% to 56%) but was approximately independent of SNR for discrimination (means: 49% for JL, 71% for BB, 48% for LS). To quantify how much perceptual accuracy improves in the EM condition relative to that in the FIX condition, we computed their relative efficiency. For both tasks, the relative efficiencies were similar and were nearly constant as a function of SNR. For detection, the mean relative efficiencies were 33% for JL, 36% for BB, and 38% for LS. For discrimination, the mean relative efficiencies were 35% for JL, 38% for BB, and 39% for LS.

D. Model Fits

To better understand saccadic and perceptual performance, we fit the decision accuracy data with a model based on SDT (see Section 2 and Appendix A). The model predictions [Eq. (A7)] are similar to the data and are shown as the curves through the data in Fig. 7. The fitted parameters summarize performance, facilitate comparison between saccades and perception, and help quantify task effects. The maximum-likelihood fitted values for the two model parameters, α and U , for first saccades and perception (FIX) are shown in Figs. 8 and 9.

The fits emphasize the two main task effects, which were shared by both saccades and perception. First, the slope parameter (α) was significantly lower for discrimination than for detection for all three observers (i.e., all the points are below the diagonal in Fig. 8). Second, the discrimination data were best fitted with zero effective uncertainty ($U = 0$, not shown in Fig. 9),³⁴ while the detection data were best fitted by effective uncertainties, which were significantly ($p < 0.05$) greater than zero (Fig. 9). The fits also emphasize two similarities between saccadic and perceptual (FIX) performance, which were found across both tasks. First, for the observers with approximately matched processing times, the slopes (α) for saccades were similar to, but slightly lower than, those for perception (FIX). Second, the uncertainties (U) for saccades were not significantly different from those for perception ($p > 0.05$).

In the EM condition, the accuracy of the final perceptual decision was also well predicted by the model. The slopes (α) for discrimination were 0.70, 0.84, and 0.69 and for detection were 0.78, 0.78, and 0.82 for JL, BB, and LS, respectively. As for the two other conditions, the effective uncertainty (U) for discrimination was zero for all three observers, while U for detection was 5, 2, and 2 for JL, BB, and LS, respectively, lower than those for the first saccade and for perception (FIX).

4. DISCUSSION

When asked to find a target in a visual search task, observers make saccadic eye movements to foveate display regions of interest. Each saccade requires a neural decision to choose when to make the saccade and what image region to foveate next. Our experiments sought to compare the visual processing used to make saccadic decisions with that used to make perceptual decisions. We hypothesized that if similar neural processing mediates both decisions, and perception and saccades are presented with equal visual information, then (1) the accuracy of perceptual and saccadic decisions should be similar, (2) the type of task should similarly affect both decisions, and (3) target salience (SNR) should modulate saccadic and perceptual performance in a similar manner.

Our results show that when retinal eccentricity and processing times are properly matched, the accuracy of the first saccadic decision is similar to that of perception for a wide range of target saliences. In the detection task, the efficiency of the perceptual (FIX) and first saccadic decisions are indistinguishable, while in the discrimination task, the first saccade is slightly less efficient. Saccadic and perceptual performance also show parallel

task effects, most notably increased effective uncertainty for detection. Furthermore, both saccadic and perceptual accuracies are well fitted by an SDT model using similar fitted parameters. Last, as expected, for both tasks, when observers were allowed to make saccadic eye movements and had longer search durations, their final perceptual decisions were much more accurate.

A. Caveats

1. Mismatches in Perceptual and Saccadic Processing Times

The two observers BB and JL, who had median saccadic latencies that correspond to approximately equal saccadic and perceptual processing times, also had similar perceptual and saccadic performance. The third observer, LS, had considerable shorter saccadic latencies, which produced much shorter processing times for saccades than for perception in the FIX condition. This was associated with LS's saccadic accuracies being considerably lower than his perceptual accuracies. Previous visual search studies^{8,9,35} have found that increasing fixation duration improves both saccadic and perceptual accuracy.

One explanation of these performance improvements is that increasing display duration causes the effective internal noise to decrease. Similarly, physiological studies^{36–38} have found that for brief displays, the SNR of neural responses increases as display duration is increased. LS's data and the fits of the SDT model are consistent with this explanation. Because LS had approximately equal latencies for each SNR and both tasks (Fig. 4), the mismatches in processing time should be similar for all SNRs and produce equal differences in internal noise. The effect of adding equal internal noise to all the visual processing mechanisms in our SDT model is a decrease in slope. This is precisely what we find for observer LS, who had shorter saccadic processing times. In the discrimination task, LS's slope for perception was 0.60, while his slope for saccades was only 0.51. In the detection task, LS's perception and saccades had nearly equal uncertainty parameters, but his slope for perception was 0.43 while his slope for saccades was only 0.30. LS's relative efficiencies are approximately constant across SNR, which is also consistent with a constant difference in the amount of internal noise.

2. Effects of Visual Masks

We chose a high-contrast noise mask to limit perceptual processing in the FIX condition, so that saccades and perception had approximately equal processing times. The actual effect of the mask on perception is more complex. One possibility is that the mask was not entirely effective and allowed some perceptual processing to continue longer than estimated. Another possibility is that the mask was too effective and disrupted perception (backward masking), which would reduce the effective perceptual processing time. Thus mask efficacy problems could produce effects similar to mismatches in perceptual and saccadic processing times. These differences would likely be similar for each SNR and produce an overall improvement/decrement in performance. Specifically, in our SDT model, increases/decreases in each mechanism's

noise correspond directly to decreases/increases in the slope parameter in our fits. For example, the slightly larger slope parameters for perception compared with those for saccades for observers BB and JL may be due to incomplete masking, allowing for slightly longer effective perceptual processing times. It is also possible that the mask had a smaller effect on perception in the discrimination task than in the detection task, which would account for the slightly lower relative efficiencies in the discrimination task.

3. Speed–Accuracy Trade-offs

Although saccadic latencies were largely independent of SNR (Fig. 3), a detailed examination of the latency data for BB and JL shows that there is a small decrease in saccadic latency as SNR increases. These latency changes are less than 20 ms for the detection task (solid curves) but are larger for the discrimination task (dashed curves), with the highest SNR having a latency that is shorter than the lowest SNR by approximately 33 ms for BB and 49 ms for JL. An examination of the model fits to the data (Fig. 6) also shows that for BB's and JL's saccadic discrimination data, the measured accuracy for the highest SNR is slightly below the model predictions. This small but systematic deviation from the model may be explained by the decreases in saccadic processing times produced by the shorter saccadic latencies at these SNRs (i.e., a speed–accuracy trade-off). This dependence of latency on SNR is somewhat surprising given that the SNRs were randomly interleaved and that others have not observed such an effect.⁹

Our finding that, for a given SNR, saccadic decision accuracy is largely unaffected by saccadic latency (Fig. 4) may seem contradictory to the potential speed–accuracy trade-off described above and to previous findings^{8,9,35} that increasing saccadic latency (fixation duration) improves both saccadic and perceptual accuracy. This may be reconciled by considering the effect of our high level of external noise. For each SNR, the addition of external noise makes the difficulty of each trial depend on the specific instance of external noise that was added to that trial. For example, on some trials, the noise added to the target location might be similar to the target shape (template) and contrast, which would increase the response of a visual mechanism responding to the target location and make the task easier. Alternatively, on other trials, the noise added to the target location might be similar to the target shape (template) but of opposite contrast, which would decrease the response of a visual mechanism responding to the target location and make the task more difficult. It is possible that the saccadic latency depends in part on the specific noise sample presented on each trial, with “easy” trials having a shorter latency and “difficult” trials having a longer latency. If this were true, it would diminish any observable effect of saccadic latency on saccadic decision accuracy, because the short-latency trials would be “easier” and have a higher chance of being correct than the more “difficult,” long-latency trials.

B. Effects of Task on Saccades and Perception

All of our observers show clear parallel task-dependent differences in saccadic and perceptual (FIX) performance

(Figs. 5 and 6). To quantify these differences, we fitted the data to an SDT model with two free parameters, a slope (α) and an uncertainty (U). We found that all three observers had significantly lower slopes and uncertainties in the discrimination task than in the detection task for both saccades and perception (Figs. 8 and 9). The slope decreases were approximately equal for saccades and perception (JL, 29% versus 33%; BB, 22% versus 17%; LS, 35% versus 28%). The similarity of the decrease for saccades and perception suggests that it may result from a similar cause. One possible explanation is a decelerating contrast nonlinearity. In our experiments, the noise amplitude was held fixed, and the SNR was varied by changing the target contrast. While the contrast increments were approximately equal for the detection and discrimination tasks, the detection task had a zero contrast pedestal and the discrimination task had a 52.7% peak contrast pedestal. Thus a decelerating contrast nonlinearity would lower the responses to the higher contrasts more than it would to the lower contrasts, resulting in the observed slope change. A similar possibility is that there is contrast-dependent internal noise.^{15,19,29,39} Experiments that vary the external noise level and the pedestal contrast would provide further information about the reason for this decrease in slope (e.g., see Refs. 15 and 39).

In the model, the discrimination data were best fitted with the uncertainty number U equal to zero for all three observers, while each of the fits to the detection data produced nonzero uncertainties. The uncertainty parameter quantifies one aspect of the task-dependent performance differences; the discrimination data are fitted well by a line passing through zero, while the detection data are not. But this difference should not be interpreted as implying that observers monitor more detectors in the detection task than in the discrimination task. In the model, observers monitor one relevant mechanism at each possible target location and U additional, orthogonal irrelevant channels. In the detection task, for low-contrast targets (SNRs), the irrelevant channels that respond only to the noise can often produce the largest response and cause errors. In the discrimination task, the addition of orthogonal irrelevant channels would not affect performance much at all. The channels corresponding to the distractors produce large responses, which are nearly always higher than those of the additional irrelevant channels (which do not respond to the pedestal, because they are orthogonal to it). In fact, in our discrimination task, the contrast of the distractors was higher than the highest target contrast that we used in the detection task. Thus, when discriminating between the relatively high-contrast distractors and the higher-contrast target, the zero-mean, irrelevant channels have little effect.³⁴ We verified that the model predictions for the discrimination task are nearly unchanged if the number of irrelevant channels is set to be equal to that of the detection task. Thus, while the detection and discrimination tasks produce differences in performance that appear to be related to stimulus uncertainty, this effect may mainly reflect differences in the stimuli rather than differences in the number of visual mechanisms used to process the stimulus.

C. Performance Improvements Due to Active Search

Previously, we reported that, not surprisingly, allowing observers to make eye movements and extending the stimulus duration produced substantial improvements in perceptual performance in a detection task.¹ Here, we replicate these findings and extend them to a discrimination task (Fig. 6). For all observers, the absolute efficiencies of the FIX perceptual decisions are less than half as high as the absolute efficiencies of the EM perceptual decisions.

There are two major differences between the EM perception condition and the two other conditions: saccades and FIX perception.¹ First, in the EM perception condition, observers had a much longer processing time (up to ~4000 ms) than in the two other conditions (~150 ms). Second, because observers were allowed to make saccadic eye movements, they could foveate the possible target locations. The stimulus duration in the EM condition was long enough for observers to fixate most, if not all, of the target locations, while in the two other conditions, processing was based on a single central fixation, with the element locations at a fixed eccentricity. The EM data show that observers did make saccades to fixate many of the possible target locations. The number of saccades observers made on each trial depended on the SNR. Trials with high SNRs produced fewer saccades, while the difficult trials with low SNRs produced more saccades. Analysis of the scan path can be used to measure how task information accumulates over saccades.^{40,41}

Despite these eccentricity and duration differences, the EM perceptual data were also well fitted by the SDT model. The ability to saccade to target locations and foveate them is likely to have two effects on performance. First, the smaller foveal receptive fields and the higher receptive field density make it likely that observers would be able to utilize visual processing mechanisms, with receptive field shapes better matched to the target spatial profile.⁴² This suggests that the slope parameter values for the EM perception condition should be higher than those for the FIX perception condition. Indeed, all observers showed large increases in the slope parameter for both tasks (mean increases were 29% for detection and 59% for discrimination). Second, observers would be likely to have better information about the location and the shape of the target and monitor fewer irrelevant mechanisms, which would reduce the effective uncertainty parameter for the detection data. Indeed, for the detection task, U for EM perception was lower than that for FIX perception (mean: 3.0 versus 10.1). For the discrimination task, the effective uncertainties were zero for both FIX and EM perception.

D. Absolute and Relative Efficiencies

Although it may appear obvious, the degree to which saccades are guided by visual information must depend on the salience of the visual information available. At one extreme, if little or no visual information is available, the neural decisions that control saccade generation will likely be dominated by internal noise or influenced by cognitive strategies, and the saccades will appear to be uncorrelated with visual information. On the other hand, if the visual information is highly salient, the neural deci-

sions will nearly always generate "correct" saccades to the target, and saccades will clearly be guided by visual information, but because the task is so easy, almost any visual mechanism would achieve similar performance. It is therefore impossible to determine "how" and "which" visual information is used. Thus the disparate results about which visual features such as color,^{3,11,12,26,43} shape,^{9,11,12,44} and symbology¹³ are used to guide saccades cannot be disentangled without measuring the role of salience. Hooze and Erkelens^{8,9} and we¹ showed that varying salience can change how much saccades appear to be guided by visual information, and it was suggested that target salience was the key to resolving the apparently conflicting results. To quantify how much and how efficiently saccades are guided by visual information, we used displays containing noise with known statistics and signals of varying strengths. This type of display allowed us to use SDT to measure the absolute efficiency with which the visual information is being used and to directly compare performance in different tasks and at different signal strengths (salience) by computing efficiencies.

For the detection task, the absolute efficiencies of the three performance measures all increase as SNR increases. As discussed in Section 3, this appears to be due to uncertainty effects, which have a larger effect on efficiency for the low SNRs. Thus the detection absolute efficiencies are lower than those for discrimination for the lowest SNRs but are similar for the highest SNRs. The absolute efficiency of the first saccade ranged between 7% and 29%, which is comparable with the previously measured efficiencies ranging from 4% to 20% in a similar task.¹ The relative efficiency of saccades to FIX perception is nearly independent of SNR for all three observers and is ~1.0 for the two observers with matched processing times, BB and JL. As discussed below, this similarity in performance provides evidence, but not proof, that the two share a performance-limiting, visual processing mechanism.

For the discrimination task, absolute efficiencies are approximately constant for the different SNRs for all three performance measures. Thus, for example, for observer BB's saccadic decisions in the EM condition, while the percentage of correct decisions ranges from 39% for the lowest SNR to 82% for the highest SNR, this apparent difference in performance reflects the difference in the task difficulty and the amount of visual information available. The absolute efficiencies show that for both of these SNRs, the observer used approximately 20% of the information available to the ideal observer. The relative efficiency of saccades to FIX perception is nearly independent of SNR for all three observers and is ~83% for the two observers with matched processing times, BB and JL. While this similarity in efficiency suggests that the perceptual and saccadic decisions share similar visual processing mechanisms,⁴⁵ other interpretations are possible.

As explained in Appendix A, d' can be expressed as the ratio of the overlap of the receptive field with the target to the total noise, where the total noise is the sum of a fixed external noise term and an internal noise source. Thus any process for which this ratio remains fixed would produce the same performance. For example, if perception

were performed by mechanisms with receptive fields very well matched to the target and high internal noise levels, and saccades were controlled by mechanisms with large receptive fields poorly matched to the target and low internal noise levels, then, although the mechanisms would be very different, the performances could still be identical. Thus, although our results are consistent with perception and saccades sharing similar visual processing mechanisms, with saccadic processing containing a small additional noise source to account for its slightly lower efficiency, it remains possible that the visual processing for each has different receptive fields and internal noise sources and that the similar performance occurs by chance. Measuring performance for a number of external noise levels would help resolve this issue.⁴⁶

It is interesting that the relative efficiency of saccades to perception is higher for the detection task than for the discrimination task for all three observers. One possible explanation is that, for the saccadic system, it is most important to detect and localize objects and less important to discriminate between them. To examine this issue further, we have also measured perceptual and saccadic performance in a “higher-order” search task: letter discrimination.⁴⁷

5. CONCLUSION

In two types of visual search tasks, contrast detection and discrimination, we have measured human performance in three ways: (1) perceptual accuracy with long display duration and free eye movements, (2) decision accuracy of the first saccade, and (3) perceptual decision accuracy with short display durations and central fixation. We have shown that each of these performance metrics can be well fitted by an SDT search model with two free parameters: a slope and an uncertainty. The detection data required a nonzero uncertainty parameter, while all the discrimination data were fitted best with zero uncertainty. The two observers for which the processing times were well matched had detection accuracies for perception and saccades that were indistinguishable and had discrimination accuracies that were nearly identical. Finally, all observers showed similar task-dependent and salience-dependent trends for both perception and saccades. These results show that measurements of the accuracy of the first saccade provide much insight about the observer’s perceptual state at the time of the saccadic decision. These results provide evidence that saccades and perception share a similar visual processing mechanism, perhaps in the superior colliculus,^{27,48–50} the frontal eye fields,^{51,52} the lateral intraparietal cortex,^{53–55} or the inferior temporal cortex.⁵⁶

APPENDIX A

SDT models⁴ have been successfully used to predict human perceptual ability to localize targets in visual search experiments for tasks with white-noise backgrounds⁵ and more realistic backgrounds.^{57,58} They can accurately predict human detection of a target among a set of distractors differing from the target along a single physical attribute¹⁷ or many.^{18,59} Here, we describe our imple-

mentation of a SDT model for a visual search task in which we assume that there are N possible nonoverlapping target locations.

The stimulus is specified by the contrast at each pixel. Because the possible target locations are far apart, for simplicity we will denote the contrast values at a pixel (x, y) centered on target location k by $c_k(x, y)$. The stimulus is created by the following steps:

1. Generating a noise background by randomly choosing the contrast of each pixel $n_k(x, y)$ from independent Gaussian distributions with identical standard deviation σ_E and zero mean.
2. Adding the target (or the distractor) contrast to the noise background. The target contrast at location (x, y) is $A_T t(x, y)$, where A_T is a constant scaling the overall contrast, which depends on the SNR. Because the distractor spatial profile is identical to the target, the distractor contrast is $A_D t(x, y)$ ($A_D = 0$ for the detection task and is constant for the discrimination task). The coefficients $t(x, y)$ (or target template) describe the spatial profile of the target (and distractors) and are normalized so that $\sum_{x,y} [t(x, y)]^2 = 1$.

1. Ideal Observer

The ideal observer optimally uses the image data to achieve the best possible performance, the highest localization accuracy. Peterson *et al.*⁶⁰ have shown that under these stimulus conditions, when the target appears with equal probability at each of the N possible locations, the ideal observer computes the correlation R_i between a filter that exactly matches the target’s spatial profile $t(x, y)$ and the image contrast values $c_i(x, y)$ at each of the i possible signal locations, and then chooses the location i that has the largest response R_i :

$$R_i = \sum_{x,y} t(x, y) c_i(x, y). \quad (\text{A1})$$

It can be shown that all of the responses R_i are Gaussian-distributed random variables with standard deviation σ_E . The responses corresponding to distractors have means A_D , while the response corresponding to the target has mean A_T .

For these assumptions, Green and Swets⁴ showed that the proportion of correct localization decisions depends only on the difference of the means divided by their common standard deviation, and they have named this quantity d' . Because d' of the ideal observer is the ratio of the signal to the noise in the stimulus, it is also referred to as the SNR:

$$d'_I = \frac{\mu_T - \mu_D}{\sigma_R} = \frac{A_T - A_D}{\sigma_E} = \text{SNR}, \quad (\text{A2})$$

where μ_T is the mean response of a mechanism to the signal, μ_D is the mean response of a mechanism to the distractors, and σ_R is the standard deviation of the response of the template. The probability of a correct decision (P_c) is equal to the probability of the response of the mechanism corresponding to the target being greater than the responses of the $N - 1$ mechanisms corresponding to the distractors:

$$P_C(d', N) = \int_{-\infty}^{+\infty} dx P(R_T = x) P(\text{all } R_D < x), \quad (\text{A3})$$

$$P_C(d', N) = 100(2\pi)^{-N/2} \int_{-\infty}^{+\infty} dx \exp\left[-\frac{(x - d')^2}{2}\right] \times \left[\int_{-\infty}^x dy \exp\left(-\frac{y^2}{2}\right) \right]^{N-1}. \quad (\text{A4})$$

Equation (A4) is used to convert decision accuracy in an N -AFC task from percent correct units to d' units.

2. Modeling Human Performance

We assume that humans make their decision by examining the output of visual mechanisms and choosing the location with the largest response. We assume that humans use mechanisms with filters (receptive fields) $f(x, y)$ that may be not perfectly matched to the target's spatial profile $t(x, y)$ and that the visual processing mechanisms have an additional internal noise source, which represents variability in firing rate and decision criteria. Specifically, we assume the following:

1. The filter coefficients $f(x, y)$ describe the mechanism's response to the contrast at each pixel. The coefficients are normalized to unity, so that $\sum_{x,y} [f(xy)]^2 = 1$.
2. The response of each mechanism, R_i , is calculated by summing the responses to each pixel and adding internal noise N_i : $R_i = N_i + \sum_{x,y} f(x, y) c_i(x, y)$.
3. The internal noise values N_i are Gaussian distributed with zero mean and standard deviation σ_I and are independent of each other.
4. The response of each mechanism to the stimulus can be shown to be a Gaussian-distributed random number.
 - a. Its mean is determined by the stimulus strength at its location, A_i , and the overlap of the receptive field with the target template: $\mu_i = A_i \sum_{x,y} f(x, y) t(x, y)$.
 - b. Its standard deviation is determined by the internal and external noise levels. The external noise standard deviation is equal to the standard deviation of each pixel, σ_E , because the filter is normalized to unity. The total noise is the sum of the internal and external noise sources: $\sigma_R^2 = \sigma_I^2 + \sigma_E^2$.

To fit the discrimination data, we assume that there are exactly N visual mechanisms, one centered on each possible target location. Thus there are $N - 1$ mechanisms that respond to distractors, each with a mean equal to $\mu_D = A_D \sum_{x,y} f(x, y) t(x, y)$, and there is a single mechanism responding to the target with mean $\mu_T = A_T \sum_{x,y} f(x, y) t(x, y)$. Then

$$d'_H = \frac{(A_T - A_D) \sum_{x,y} f(x, y) t(x, y)}{\sqrt{\sigma_E^2 + \sigma_I^2}} = \text{SNR} \times \frac{\sum_{x,y} f(x, y) t(x, y)}{[1 + (\sigma_I^2/\sigma_E^2)]^{1/2}}. \quad (\text{A5})$$

Thus human performance measured in d' units is a constant times the SNR. We call this constant the slope, α , and determine its value by fitting the data:

$$\alpha = \frac{\sum_{x,y} f(x, y) t(x, y)}{[1 + (\sigma_I^2/\sigma_E^2)]^{1/2}}. \quad (\text{A6})$$

Because the detection data were not fitted well by this simple model, we added another source of inefficiency, stimulus uncertainty. As Pelli^{16,30} suggested, humans often perform as if they were uncertain of the target location and possibly the target shape. To model this, we assumed that instead of monitoring a single visual mechanism at each target location, humans monitored U additional mechanisms per location. We again assume that the decision rule is to choose the location corresponding to the mechanism with the highest response. To make the calculations tractable, we assumed the following:

1. The mechanisms' responses were independent:

$$\sum_{x,y} f_i(x, y) f_j(x, y) = 0 \text{ for } i \neq j.$$

2. Only one mechanism at each location responded to the signal/distractor, while the others responded to noise:

$$\sum_{x,y} f_i(x, y) t(x, y) = 0 \text{ for all except one mechanism.}$$

3. The standard deviations of the responses were equal.

While it is possible to relax these assumptions and calculate the predicted performance, doing so adds additional free parameters. An observer monitoring $U + 1$ mechanisms per location will be correct if any of the mechanisms corresponding to the target location has the highest response. For our detection task, the distractors had zero contrast, so the responses of all mechanisms except that responding to the target had equal means (zero). As shown by Eckstein *et al.*,¹⁹ in this case, the percent correct is

$$P_C(d', N) = 100(2\pi)^{-N(U+1)/2} \int_{-\infty}^{+\infty} dx \exp\left[-\frac{(x - d')^2}{2}\right] \times \left[\int_{-\infty}^x dy \exp\left(-\frac{y^2}{2}\right) \right]^{N(U+1)-1} + 100U(2\pi)^{-N(U+1)/2} \int_{-\infty}^{+\infty} dx \exp\left(-\frac{x^2}{2}\right) \times \left[\int_{-\infty}^x dy \exp\left(-\frac{y^2}{2}\right) \right]^{N(U+1)-2} \times \int_{-\infty}^x dz \exp\left[-\frac{(z - d')^2}{2}\right], \quad (\text{A7})$$

where, as above,

$$d' = \text{SNR} \times \frac{\sum_{x,y} f(x,y)t(x,y)}{[1 + (\sigma_i^2/\sigma_E^2)]^{1/2}}$$

$$= \alpha \times \text{SNR}.$$

We fit this model to the detection data by using the two free parameters: the uncertainty number, U , and the slope, α .

3. Comparing Performance

A common metric to quantify how well humans perform a task is to compare their performance with that of the ideal observer with a metric known as the absolute efficiency.^{5,33} We define absolute efficiency⁶¹ as

$$\text{Absolute efficiency} = (d'_{\text{human}}/d'_{\text{ideal}})^2. \quad (\text{A8})$$

Thus, for our model with zero uncertainty, the absolute efficiency measures how well the visual processing filter matches the signal template and how large the internal noise is relative to the external noise:

$$\text{Absolute efficiency} = \left(\frac{\sum_{x,y} f(x,y)t(x,y)}{[1 + (\sigma_i^2/\sigma_E^2)]^{1/2}} \right)^2. \quad (\text{A9})$$

For the model with stimulus uncertainty, the absolute efficiency depends on SNR. To compare performance, a useful metric is the relative efficiency, which measures the amount of information used by one decision relative to that used by another:

$$\text{Relative efficiency} = (d'_{\text{decision 1}}/d'_{\text{decision 2}})^2. \quad (\text{A10})$$

ACKNOWLEDGMENTS

This research was supported by NASA grant NAG-1157 to Miguel P. Eckstein and by NASA Airspace Systems (727-05-30) and Biomedical Research and Countermeasures (111-10-10) Program support to Leland S. Stone.

Address correspondence to Brent R. Beutter at the location on the title page or by phone, 650-604-5150; fax, 650-604-0255; or e-mail, bbeutter@mail.arc.nasa.gov.

REFERENCES AND NOTES

1. M. P. Eckstein, B. R. Beutter, and L. S. Stone, "Quantifying the performance limits of human saccadic targeting during visual search," *Perception* **30**, 1389–1401 (2001).
2. B. R. Beutter, M. P. Eckstein, and L. S. Stone, "Parallel differences in contrast-discrimination and detection performance for saccades and perception in visual search," *Invest. Ophthalmol. Visual Sci. (Suppl.)* **41**, S424 (2000).
3. J. M. Findlay, "Saccade target selection during visual search," *Vision Res.* **37**, 617–631 (1997).
4. D. M. Green and J. A. Swets, *Signal Detection Theory and Psychophysics* (Wiley, New York, 1966).
5. A. E. Burgess, R. F. Wagner, R. J. Jennings, and H. B. Barlow, "Efficiency of human visual signal discrimination," *Science* **214**, 93–94 (1981).
6. P. Viviani and R. G. Swensson, "Saccadic eye movements to peripherally discriminated visual targets," *J. Exp. Psychol. Hum. Percept. Perform.* **16**, 459–478 (1982).
7. C. Motter and E. J. Belky, "The guidance of eye movements during active visual search," *Vision Res.* **38**, 1805–1815 (1998).
8. T. Hooge and C. J. Erkelens, "Adjustment of fixation duration in visual search," *Vision Res.* **38**, 1295–1302 (1998).
9. T. Hooge and C. J. Erkelens, "Peripheral vision and oculomotor control during visual search," *Vision Res.* **39**, 1567–1575 (1999).
10. J. M. Findlay, V. A. Brown, and I. D. Gilchrist, "Saccade target selection in visual search: the influence of information from the previous fixation," *Vision Res.* **41**, 87–95 (2001).
11. L. G. Williams, "The effects of target specification on objects fixated during visual search," *Acta Psychol.* **27**, 355–360 (1966).
12. L. G. Williams, "Target conspicuity and visual search," *Hum. Factors* **8**, 80–92 (1967).
13. G. L. Zelinsky, "Using eye saccades to assess the selectivity of search movements," *Vision Res.* **36**, 2177–2187 (1996).
14. G. E. Legge and J. M. Foley, "Contrast masking in human vision," *J. Opt. Soc. Am.* **70**, 1458–1471 (1980).
15. G. E. Legge, D. Kersten, and A. E. Burgess, "Contrast discrimination in noise," *J. Opt. Soc. Am. A* **4**, 391–404 (1987).
16. D. G. Pelli, "Uncertainty explains many aspects of visual contrast detection and discrimination," *J. Opt. Soc. Am. A* **2**, 1508–1532 (1985).
17. J. Palmer, "Set-size effects in visual search: the effect of attention is independent of the stimulus for simple tasks," *Vision Res.* **34**, 1703–1721 (1994).
18. J. Palmer, P. Verghese, and M. Pavel, "The psychophysics of visual search," *Vision Res.* **40**, 1227–1268 (2000).
19. M. P. Eckstein, A. J. Ahumada, and A. B. Watson, "Visual signal detection in structured backgrounds. II. Effect of contrast gain control, background variations and white noise," *J. Opt. Soc. Am. A* **14**, 2406–2419 (1997).
20. The SNRs are slightly different for the discrimination and detection stimuli because we measured each SNR from the stimuli actually used in the experiment rather than relying on the ensemble parameters used to generate the stimuli.
21. W. Becker and R. Jürgens, "An analysis of the saccadic system by means of double step stimuli," *Vision Res.* **19**, 967–983 (1979).
22. I. T. Hooge, "Control of eye movement in visual search," Ph.D. thesis (Utrecht University, Utrecht, The Netherlands, 1996).
23. P. He and E. Kowler, "The role of location probability in the programming of saccades: implications for center-of-gravity tendencies," *Vision Res.* **29**, 1165–1181 (1989).
24. L. S. Stone, B. R. Beutter, and M. P. Eckstein, "Salience effects on perceptual and saccadic target localization during search," *Soc. Neurosci. Abstr.* **25**, 548 (1999).
25. B. R. Beutter and L. S. Stone, "Human motion perception and smooth eye movements show similar directional biases for elongated apertures," *Vision Res.* **38**, 1273–1286 (1998).
26. R. M. McPeck, A. A. Skavenski, and K. Nakayama, "Concurrent processing of saccades in visual search," *Vision Res.* **40**, 2499–2516 (2000).
27. R. M. McPeck and E. L. Keller, "Superior colliculus activity related to concurrent processing of saccade goals in a visual search task," *J. Neurophysiol.* **87**, 1805–1815 (2002).
28. An alternative model has been proposed by Lu and Doshier²⁹ in which a nonlinear transducer replaces the intrinsic uncertainty.
29. Z. L. Lu and B. A. Doshier, "Characterizing human perceptual inefficiencies with equivalent internal noise," *J. Opt. Soc. Am. A* **16**, 764–778 (1999).
30. D. G. Pelli, "Effects of visual noise," Ph.D. thesis (Cambridge University, Cambridge, UK, 1981).
31. This decision strategy is suboptimal, as discussed in Pelli,¹⁶ but approximates the ideal decision at high SNRs,³² although for the parameters that we use the difference in performance is small.
32. L. W. Nolte, and D. Jaarsma, "More on the detection of one of M orthogonal signals," *J. Acoust. Soc. Am.* **41**, 497–505 (1967).
33. H. B. Barlow, "The absolute efficiency of perceptual decisions," *Proc. R. Soc. London* **290**, 71–91 (1980).

34. One should note that when applying the uncertainty equation¹⁹ to a contrast-discrimination task, one should interpret U as the effect of uncertainty on performance rather than the number of statistically independent signal-irrelevant responses monitored by the observer. In addition to the discriminability (d_1) of the signal with respect to the distractors, an alternative model for the contrast-discrimination task would take into account the discriminability (d_2) of the signal with respect to the U additional signal-irrelevant mechanisms. This more complete formulation is given by

$$P_C = 100 \int_{-\infty}^{+\infty} dx [g(x)G(x + d_1')^{N-1}G(x + d_2')^{UN} + Ug(x + d_2')G(x)G(x + d_1')^{N-1}G(x + d_2')^{UN1}],$$

where $g(x)$ is the Gaussian density function, $G(x)$ is the cumulative probability, U is the number of additional signal-irrelevant mechanisms per location monitored, and N is the number of possible target locations. This model has two fitting parameters and would require a separate experiment to estimate d_2' . Note that if d_2' is large, as in the contrast-discrimination task (>2.5 approximately), the irrelevant mechanisms almost never produce the largest response, and increasing U in the equation above has very little effect on P_C .

35. W. S. Geisler and L. Chou, "Separation of low-level and high-level factors in complex tasks: visual search," *Psychol. Rev.* **102**, 356–378 (1995).
36. D. J. Tolhurst, J. A. Movshon, and F. A. Dean, "The statistical reliability of signals in single neurons in cat and monkey visual cortex," *Vision Res.* **23**, 775–785 (1983).
37. W. S. Geisler and D. G. Albrecht, "Bayesian analysis of identification performance in monkey visual cortex: nonlinear mechanisms and stimulus certainty," *Vision Res.* **35**, 2723–2730 (1995).
38. W. S. Geisler and D. G. Albrecht, "Visual cortex neurons in monkeys and cats: detection, discrimination and identification," *Visual Neurosci.* **14**, 897–919 (1997).
39. A. E. Burgess and B. Colborne, "Visual signal detection. IV. Observer inconsistency," *J. Opt. Soc. Am. A* **5**, 617–627 (1988).
40. M. P. Eckstein, B. R. Beutter, and L. S. Stone, "Accumulation of information across saccades during visual search depends on how far the first saccade lands from the target," *Perception (Suppl.)* **29** (2000), <http://www.perceptionweb.com/perception/ecvp00/0029.html>.
41. M. P. Eckstein, B. R. Beutter, and L. S. Stone, "Task information increases from the first to the second saccade in visual search of a target among distractors," *Invest. Ophthalmol. Visual Sci.* **41**, 759 (2000).
42. R. F. Hess and A. Hayes, "The coding of spatial position by the human visual system: effects of spatial scale and retinal eccentricity," *Vision Res.* **34**, 625–643 (1994).
43. R. M. McPeck, V. Maljkovic, and K. Nakayama, "Saccades require focal attention and are facilitated by a short-term memory system," *Vision Res.* **39**, 1555–1566 (1999).
44. D. Gilchrist, C. A. Heywood, and J. M. Findlay, "Saccade selection in visual search: evidence for spatial frequency specific between-item interactions," *Vision Res.* **39**, 1373–1383 (1999).
45. B. R. Beutter, L. S. Stone, and M. P. Eckstein, "Correlated saccadic and perceptual decisions in a visual-search detection task reveal spatial-filter overlap," presented at the Vision Sciences Society Meeting, May 4–8, 2001, Sarasota, Fla., *J. Vision* **1**, No. 1 (Abstract 263) (2001), <http://www.journalofvision.org/1/3/263/>.
46. B. R. Beutter, M. P. Eckstein, and L. S. Stone, "Similar internal noise levels limit saccadic and perceptual performance in a visual-search task," Program No. 418.13 (2002). Abstract Viewer/Itinerary Planner. Society for Neuroscience, Washington, D.C., 2002. Online. <http://sfn.scholarone.com/itin2002/index.html>.
47. R. F. Murray, B. R. Beutter, M. P. Eckstein, and L. S. Stone, "Saccadic and perceptual performance in visual search tasks. II. Letter discrimination" *J. Opt. Soc. Am. A* **20**, 1356–1370 (2003).
48. R. J. Krauzlis, A. Z. Zivotofsky, and F. A. Miles, "Target selection for pursuit and saccadic eye movements in humans," *J. Cogn. Neurosci.* **11**, 641–649 (1999).
49. M. A. Basso and R. H. Wurtz, "Modulation of neuronal activity by target uncertainty," *Nature (London)* **389**, 66–69 (1997).
50. M. A. Basso and R. H. Wurtz, "Modulation of neuronal activity in superior colliculus by changes in target probability," *J. Neurosci.* **18**, 7519–7534 (1998).
51. J. D. Schall, "Neural basis of saccade target selection," *Rev. Neurosci.* **6**, 63–85 (1995).
52. J. D. Schall and K. G. Thompson, "Neural selection and control of visually guided eye movements," *Annu. Rev. Neurosci.* **22**, 241–259 (1999).
53. J. P. Gottlieb, M. Kusunoki, and M. E. Goldberg, "The representation of visual salience in monkey parietal cortex," *Nature (London)* **391**, 481–484 (1998).
54. C. L. Colby and M. E. Goldberg, "Space and attention in parietal cortex," *Annu. Rev. Neurosci.* **23**, 319–349 (1999).
55. P. W. Glimcher, "Making choices: the neurophysiology of visual-saccadic decision making," *Trends Neurosci.* **24**, 654–659 (2001).
56. L. Chelazzi, E. K. Miller, J. Duncan, and R. Desimone, "A neural basis for visual search in inferior temporal cortex," *Nature (London)* **363**, 345–347 (1993).
57. M. P. Eckstein and J. S. Whiting, "Visual signal detection in structured backgrounds. I. Effect of number of possible signal locations and signal contrast," *J. Opt. Soc. Am. A* **13**, 1777–1787 (1996).
58. M. P. Eckstein, J. S. Whiting, and J. P. Thomas, "Role of knowledge in human visual temporal integration in spatiotemporal noise," *J. Opt. Soc. Am. A* **13**, 1960–1968 (1996).
59. M. P. Eckstein, "The lower efficiency for conjunctions is due to noise and not serial visual attention," *Psychol. Sci.* **9**, 111–118 (1998).
60. W. W. Peterson, T. G. Birdsall, and W. C. Fox, "The theory of signal detectability," *IRE Trans. Inf. Theory* **PGIT-4**, 171–212 (1954).
61. This definition compares human- and ideal-observer performances at the same SNR. As suggested by a reviewer, it is also possible to define efficiency as a comparison between the human- and ideal-observer SNRs required to achieve the same performance level. In this alternative definition, efficiency is equal to $(\text{SNR}_{\text{ideal}}/\text{SNR}_{\text{human}})^2$. The two definitions are equivalent if human d' is directly proportional to SNR (uncertainty is equal to zero), as is the case for our discrimination task. They are not the same for nonzero uncertainty (nonzero intercept), as is the case for our detection task.